**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I INFORMACIJSKIH TEHNOLOGIJA**

**Sveučilišni studij**

**POBOLJŠANJE REZULTATA SEGMENTACIJE ODJEĆE**

**Obrada slike i računalni vid**

**Kristian Rücker**

**Osijek, 2024.**

SADRŽAJ

[1. UVOD 1](#_Toc177517157)

[2. SEMATIČKA SEGMENTACIJA 2](#_Toc177517158)

[3. TEHNIKE OBRADE SLIKE ZA POBOLJŠANJE SEGMENTACIJSKIH MAPA 3](#_Toc177517159)

[4. MORFOLOŠKI OPERATORI 4](#_Toc177517160)

[5. IZRADA PROJEKTNOG ZADATKA 8](#_Toc177517161)

[5.1. Preuzimanje podatkovnog skupa 8](#_Toc177517162)

[Kod 5.1: Učitavanja originalnih slika i njihovih maski sa lokalnog računala 8](#_Toc177517163)

[5.2. Spremanje izlaza mreže sa svaku sliku 9](#_Toc177517164)

[Kod 5.2: Spremanje predikcija 9](#_Toc177517165)

[5.3. Implementacija postprocesiranja 10](#_Toc177517166)

[Kod 5.3: Korištenje morfoloških operatora i spremanje postpredikcija 10](#_Toc177517167)

[6. EVALUACIJA UČINKOVITOSTI 11](#_Toc177517168)

[6.1. Dice koeficijent 11](#_Toc177517169)

[6.2. IoU koeficijent 11](#_Toc177517170)

[Kod 5.4: Korištenje morfoloških operatora i spremanje postpredikcija 12](#_Toc177517171)

[7. ZAKLJUČAK 13](#_Toc177517172)

[LITERATURA 14](#_Toc177517173)

# UVOD

Segmentacija slika na semantičkoj razini predstavlja proces klasifikacije svakog pojedinog piksela u određenoj slici u jednu od više klasa objekata. Za razliku od klasične klasifikacije slika, koja određuje kojoj klasi pripada cijela slika, semantička segmentacija pruža detaljne informacije o prisutnosti i položaju različitih objekata unutar slike. Primjerice, ako imamo sliku na kojoj se nalazi zebra okružena travom, stablom i letećom pticom, klasična klasifikacija bi samo identificirala sliku kao "zebru". Nasuprot tome, semantička segmentacija omogućuje nam da točno odredimo koji pikseli pripadaju zebri, travi, stablu i ptici.

U ovom radu koristit ćemo gotovu neuronsku mrežu za segmentaciju odjeće na zadanim slikama iz podatkovnog skupa. Cilj je poboljšati rezultate koje mreža daje primjenom standardnih tehnika obrade slike, poput morfoloških operatora, uvjetnih slučajnih polja i drugih metoda. Istražit ćemo i testirati različite pristupe te ih primijeniti na predikcijama modela za testni skup podataka. Nakon toga, izračunat ćemo često korištena mjerila za evaluaciju segmentacije kako bismo kvantitativno i kvalitativno usporedili kvalitetu segmentacije prije i nakon obrade.

# SEMATIČKA SEGMENTACIJA

Semantička segmentacija predstavlja proces klasifikacije svakog piksela u slici u jednu od predefiniranih klasa, omogućavajući detaljno razumijevanje sadržaja slike [1]. Ova tehnika omogućuje preciznu identifikaciju i lokalizaciju objekata unutar slike, što je ključno za mnoge primjene u računalnom vidu kao što su autonomna vozila, medicinska dijagnostika i robotska percepcija [2].

Razvoj dubokih neuronskih mreža značajno je unaprijedio performanse semantičke segmentacije. Pionirski radovi kao što su Potpuno Konvolucijske Mreže (Fully Convolutional Networks - FCN) [3] omogućili su kraj-do-kraj učenje za segmentaciju slika. FCN arhitektura zamjenjuje potpuno povezane slojeve s konvolucijskim slojevima, omogućavajući mreži da generira segmentacijsku mapu koja odgovara dimenzijama ulazne slike.

Kasnije su uvedene naprednije arhitekture poput U-Net [4], koja koristi skip-veze za kombiniranje nisko-razinskih i visoko-razinskih značajki, poboljšavajući tako preciznost segmentacije. DeepLab serija modela [5] uvela je atrous konvolucije i potpuno povezane uvjetne slučajne mreže (CRF) za poboljšanje segmentacije objekata različitih veličina.

Usprkos značajnim napretcima, izazovi poput varijabilnosti osvjetljenja, pozadinske složenosti i ograničenih podataka za učenje još uvijek otežavaju postizanje savršene segmentacije [2]. Stoga se često kombiniraju metode dubokog učenja s tradicionalnim tehnikama obrade slike kako bi se poboljšali rezultati.

# TEHNIKE OBRADE SLIKE ZA POBOLJŠANJE SEGMENTACIJSKIH MAPA

Iako duboke neuronske mreže pružaju moćne alate za semantičku segmentaciju, rezultirajuće segmentacijske mape često zahtijevaju dodatnu obradu kako bi se uklonili nedostaci poput šuma, netočnih rubova i malih izoliranih područja [6]. Tehnike obrade slike igraju ključnu ulogu u poboljšanju ovih rezultata, omogućavajući rafiniranje segmentacijskih mapa za precizniju analizu.

Jedna od uobičajenih tehnika je primjena morfoloških operatora koji omogućavaju manipulaciju geometrijskih struktura unutar slike [7]. Operacije poput erozije, dilatacije, otvaranja i zatvaranja mogu se koristiti za uklanjanje šuma, popunjavanje rupa i zaglađivanje rubova [8].

Uvjetna slučajna polja (Conditional Random Fields - CRF) također se često koriste za poboljšanje segmentacijskih rezultata [9]. CRF modeli uzimaju u obzir prostorne odnose između piksela, omogućavajući konzistentnije i prostorno glatke segmentacijske mape.

Filtriranje slike, poput medijanskog ili Gaussovog filtriranja, može pomoći u uklanjanju šuma i zaglađivanju segmentacijskih mapa [10]. Tehnike klasteriranja, poput K-sredina ili spektralnog klasteriranja, mogu se koristiti za grupiranje sličnih područja i poboljšanje homogenosti segmenata [11].

Integracija ovih tehnika s rezultatima neuronskih mreža može značajno poboljšati kvalitetu segmentacije, što je posebno korisno u aplikacijama gdje je preciznost kritična.

# MORFOLOŠKI OPERATORI

Morfološki operatori su skup alata u matematičkoj morfologiji koji se koriste za obradu i analizu geometrijskih struktura u digitalnim slikama [7]. Oni omogućuju manipulaciju oblikom i strukturom objekata unutar slike putem operacija temeljenih na teoriji skupova.

Osnovni morfološki operatori uključuju:

* **Erozija**: Proces koji smanjuje objekte u slici uklanjanjem piksela na njihovim rubovima. Erozija se koristi za uklanjanje sitnih objekata ili šuma te za odvajanje spojenih objekata [12], te je prikazan na slici 4.1 [13].

**A white letter on a black background

Description automatically generated**A white letter on a black background

Description automatically generated  
**Slika 4.1:** Morfološki operator erozije

* **Dilatacija**: Suprotno eroziji, dilatacija povećava objekte dodavanjem piksela na njihove rubove. Koristi se za popunjavanje malih rupa unutar objekata i spajanje bliskih objekata [12], te je prikazan na slici 4.2 [13].

**A white letter on a black background

Description automatically generated**A white letter on a black background

Description automatically generated  
**Slika 4.2:** Morfološki operator dilatacije

* **Otvaranje**: Kombinacija erozije praćene dilatacijom. Otvaranje uklanja male objekte ili šum bez značajnog utjecaja na veličinu većih objekata [12], te je prikazan na slici 4.3 [13].

A close-up of a logo

Description automatically generated  
**Slika 4.3:** Morfološki operator otvaranja

* **Zatvaranje**: Kombinacija dilatacije praćene erozijom. Zatvaranje popunjava male rupe unutar objekata i zaglađuje njihove rubove [12], te je prikazan na slici 4.4 [13].

A close-up of a black background

Description automatically generated  
**Slika 4.4:** Morfološki operator zatvaranja

* **Gradijent:** Morfološki gradijent je operacija koja izračunava razliku između dilatirane i erodirane verzije slike. Ova operacija naglašava rubove objekata u slici, ističući prijelaze između različitih struktura [12]. Koristi se za izdvajanje kontura objekata i poboljšanje detekcije rubova u segmentacijskim mapama], te je prikazan na slici 4.5 [13].

A white line drawing of a person

Description automatically generated  
**Slika 4.5:** Morfološki operator gradijent

* **Top-Hat transformacija:** Razlika između originalne slike i njene otvorene verzije (originalna slika minus otvaranje slike). Ova operacija omogućuje izdvajanje malih objekata ili detalja koji su svjetliji od svoje okoline i koji su manji od elementa strukturnosti korištenog u otvaranju [12]. Koristi se za poboljšanje kontrasta i isticanje svijetlih detalja u slici, te je prikazan na slici 4.6 [13].

A white letters on a black background

Description automatically generated  
**Slika 4.6:** Morfološki operator Top-Hat transformacija

* **Black-Hat transformacija:** Razlika između zatvorene verzije slike i originalne slike (zatvaranje slike minus originalna slika). Ova operacija omogućuje izdvajanje tamnih objekata ili detalja koji su tamniji od svoje okoline i koji su manji od elementa strukturnosti korištenog u zatvaranju [12]. Koristi se za isticanje tamnih detalja u slici koji nisu vidljivi u originalnoj slici, te je prikazan na slici 4.7 [13].

A black and white image of a person

Description automatically generated  
**Slika 4.7:** Morfološki operator Black-Hat transformacija

Primjena morfoloških operatora na segmentacijske mape može značajno poboljšati kvalitetu segmentacije. Na primjer, erozija može ukloniti izolirane piksele koji su pogrešno klasificirani, dok dilatacija može popuniti male praznine unutar segmenata [14]. Otvaranje i zatvaranje pomažu u uklanjanju šuma i zaglađivanju rubova, što rezultira čišćim i preciznijim segmentacijskim mapama [10].

Odabir odgovarajućeg strukturnog elementa (npr. oblik i veličina kernela) ključan je za uspješnu primjenu morfoloških operatora [10]. Strukturni element definira kako će operator interagirati s objektima u slici. U praksi se često koriste jednostavni oblici poput kvadrata ili kruga, ali prilagodba veličine može značajno utjecati na rezultate [14].

Kombinacija morfoloških operatora s drugim tehnikama, poput filtriranja ili klasteriranja, može dodatno poboljšati kvalitetu segmentacije [10]. Ove metode zajedno omogućuju rafiniranje segmentacijskih mapa kako bi se postigla veća točnost i pouzdanost u praktičnim primjenama.

# IZRADA PROJEKTNOG ZADATKA

U ovom poglavlju detaljno ćemo opisati postupak realizacije projektnog zadatka, uključujući preuzimanje podatkovnog skupa, spremanje izlaza mreže za svaku sliku te implementaciju postprocesiranja.

## Preuzimanje podatkovnog skupa

Za potrebe ovog rada, podatkovni skup je preuzet i pohranjen lokalno kako bismo imali pristup originalnim slikama i pripadajućim ground truth segmentacijskim mapama, kao što je vidljivo kodom 5.1. Ovo je omogućilo detaljnu analizu i usporedbu između predikcija modela i stvarnih oznaka klase za svaki piksel.

|  |
| --- |
| image\_path = []  for root, dirs, files in os.walk('IMAGES'):      # iterate over 1000 images      for file in files:          # create path          path = os.path.join(root,file)          # add path to list          image\_path.append(path)  len(image\_path)  mask\_path = []  for root, dirs, files in os.walk('MASKS'):      #iterate over 1000 masks      for file in files:          # obtain the path  path = os.path.join(root,file)          # add path to the list          mask\_path.append(path)  len(mask\_path) |

## Kod 5.1: Učitavanja originalnih slika i njihovih maski sa lokalnog računala

## Spremanje izlaza mreže sa svaku sliku

Kaggle bilježnica koja implementira gotovu neuronsku mrežu za segmentaciju odjeće prilagođena je kako bi se omogućilo spremanje izlaza mreže za svaku sliku u testnom skupu. Izlaz mreže, odnosno predikcijske segmentacijske mape, pohranjene su u odgovarajućem formatu radi daljnje obrade i analize, a to je vidljivo kodom 5.2. Ova prilagodba omogućila je izračun metrika poput Dice koeficijenta, Intersection over Union (IoU) između predikcija modela i ground truth podataka. Primjer predikcija prikazan je slikom 5.1.

|  |
| --- |
| def save\_predictions(predictions, save\_dir):      if not os.path.exists(save\_dir):          os.makedirs(save\_dir)      for i, prediction in enumerate(predictions):          prediction = (prediction \* 255).astype(np.uint8)          if prediction.shape[-1] == 1:              prediction = prediction.squeeze(axis=-1)          elif prediction.shape[-1] == 59:              prediction = prediction[:, :, 0]          img = Image.fromarray(prediction)          img.save(os.path.join(save\_dir, f'prediction\_{i}.png')) |

## Kod 5.2: Spremanje predikcija

A person with her hands on her hips

Description automatically generated  
**Slika 5.1:** Primjer predikcije

## Implementacija postprocesiranja

Nakon dobivanja predikcijskih segmentacijskih mapa, implementiran je algoritam za postprocesiranje s ciljem poboljšanja rezultata segmentacije. Kao početni korak, primijenjeni su morfološki operatori otvaranja i zatvaranja. Ove operacije omogućile su uklanjanje šuma, popunjavanje manjih rupa unutar segmenata te zaglađivanje rubova objekata. Kod korišten za implementaciju morfoloških operatora te spremanje postprocessinga prikazan je kodom 5.3. Primjer predikcija prikazan je slikom 5.2.

|  |
| --- |
| def post\_process\_image(prediction):      if prediction.dtype != np.uint8:          prediction = prediction.astype(np.uint8)      kernel = np.ones((3, 3), np.uint8)      processed\_img = cv2.morphologyEx(prediction, cv2.MORPH\_CLOSE, kernel)      processed\_img = cv2.morphologyEx(processed\_img, cv2.MORPH\_OPEN, kernel)      return processed\_img  def save\_post\_processed\_predictions(predictions, save\_dir):      global post\_pred\_img      post\_pred\_img = []      if not os.path.exists(save\_dir):          os.makedirs(save\_dir)      for i, prediction in enumerate(predictions):          img = Image.fromarray(prediction)          img.save(os.path.join(save\_dir, f'post\_processed\_prediction\_{i}.png'))          post\_pred\_img.append(img) |

## Kod 5.3: Korištenje morfoloških operatora i spremanje postpredikcija

A comparison of a person with a weight

Description automatically generated with medium confidence**Slika 5.2:** Primjer postpredikcije

# EVALUACIJA UČINKOVITOSTI

U ovome poglavlju opisat će se metrike za evaluaciju učinkovitosti morfoloških operatora. Izračunate su segmentacijske metrike kako bi se kvantitativno procijenila kvaliteta segmentacije. Posebna pažnja posvećena je usporedbi između predikcija prije i nakon postprocesiranja u odnosu na ground truth podatke. Za evaluaciju korištene su metrike Dice koeficijent i IoU. Kod za izračun tih koeficijena prikazan je kodom 5.4, a rezultati su prikazani slikom 5.3.

## Dice koeficijent

Dice koeficijent, također poznat kao Diceova sličnost ili F1 mjera, je statistička mjera koja procjenjuje sličnost između dvaju skupova podataka. U kontekstu segmentacije slike, on mjeri preklapanje između predviđenih piksela određene klase i stvarnih piksela te klase u ground truth slici. Matematički, Dice koeficijent definira se kao:

## IoU koeficijent

Koeficijent presjeka preko unije, poznat i kao Jaccardov indeks, je još jedna mjera koja kvantificira sličnost između dvaju skupova podataka. U segmentaciji slike, IoU mjeri omjer između veličine presjeka predviđenih i stvarnih piksela te klase i veličine njihove unije. Matematički, IoU definira se kao:

|  |
| --- |
| def dice\_coefficient(y\_true, y\_pred):      y\_true\_f = y\_true.flatten().astype(np.float32)      y\_pred\_f = y\_pred.flatten().astype(np.float32)      y\_true\_f = np.where(y\_true\_f > 0, 1, 0)      y\_pred\_f = np.where(y\_pred\_f > 0, 1, 0)        intersection = np.sum(y\_true\_f \* y\_pred\_f)      union = np.sum(y\_true\_f) + np.sum(y\_pred\_f)      return (2.0 \* intersection) / (union + 1e-7)  def jaccard\_index(y\_true, y\_pred):      y\_true\_f = y\_true.flatten().astype(np.float32)      y\_pred\_f = y\_pred.flatten().astype(np.float32)      y\_true\_f = np.where(y\_true\_f > 0, 1, 0)      y\_pred\_f = np.where(y\_pred\_f > 0, 1, 0)        intersection = np.sum(y\_true\_f \* y\_pred\_f)      union = np.sum(y\_true\_f) + np.sum(y\_pred\_f) - intersection        return intersection / (union + 1e-7) |

## Kod 5.4: Korištenje morfoloških operatora

A screenshot of a computer program

Description automatically generated**Slika 5.3:** Rezultati Dice i IoU koeficijenta

# ZAKLJUČAK

U ovom radu istražena je primjena standardnih tehnika obrade slike, konkretno morfoloških operatora otvaranja i zatvaranja, za poboljšanje segmentacijskih mapa dobivenih iz neuronske mreže za segmentaciju odjeće. Nakon preuzimanja i pripreme podatkovnog skupa, prilagođena je postojeća Kaggle bilježnica kako bi se omogućilo spremanje izlaza mreže za svaku sliku te izračunavanje metrika poput Dice koeficijenta i koeficijenta presjeka preko unije (IoU) između predikcija i ground truth podataka.

Primjenom morfoloških operatora otvaranja i zatvaranja na predikcijske segmentacijske mape, uklonjen je šum, popunjene su manje rupe unutar segmenata te su zaglađeni rubovi objekata. Vizualna analiza pokazala je da su glavne razlike između predikcija i ground truth slika bile prisutnost malih izoliranih područja i nepravilnosti na rubovima segmenata. Primjena ovih morfoloških operatora učinkovito je ispravila navedene nedostatke.

Kvantitativna evaluacija rezultata pokazala je da postprocesiranje segmentacijskih mapa pomoću morfoloških otvaranja i zatvaranja poboljšava vrijednosti Dice i IoU koeficijenata u usporedbi s originalnim predikcijama iz neuronske mreže. Ovo poboljšanje ukazuje na učinkovitost korištenih tehnika obrade slike u povećanju preciznosti segmentacije.

Zaključno, dokazano je da primjena jednostavnih morfoloških operatora može značajno poboljšati rezultate semantičke segmentacije dobivene iz dubokih neuronskih mreža. Budući radovi mogli bi uključivati istraživanje drugih morfoloških operatora i tehnika obrade slike, poput uvjetnih slučajnih polja ili naprednih filtriranja, kao i njihovu primjenu na različite podatkovne skupove i klase objekata.

# LITERATURA

1. Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Rodriguez, J. G. (2018). A review on deep learning techniques applied to semantic segmentation. *Applied Soft Computing*, 70, 41-65.
2. Chen, L. C., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, dostupan na <https://arxiv.org/abs/1706.05587>
3. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3431-3440, dostupan na: <https://arxiv.org/abs/1411.4038>
4. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234-241, dostupan na: <https://arxiv.org/abs/1505.04597>
5. Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 834-848, dostupan na: <https://arxiv.org/abs/1606.00915>
6. Postels, J., et al. (2019). Sampling-free epistemic uncertainty estimation using approximated variance propagation. *Proceedings of the IEEE International Conference on Computer Vision*, 2931-2940.
7. Soille, P. (2003). *Morphological Image Analysis: Principles and Applications*. Springer, dostupan na: <https://link.springer.com/book/10.1007/978-3-662-05088-0>
8. Haralick, R. M., Sternberg, S. R., & Zhuang, X. (1987). Image analysis using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 532-550,dostupan na: <https://ieeexplore.ieee.org/document/4767941>
9. Krähenbühl, P., & Koltun, V. (2011). Efficient inference in fully connected CRFs with Gaussian edge potentials. *Advances in Neural Information Processing Systems*, 24.
10. Gonzalez, R. C., & Woods, R. E. (2002). *Digital Image Processing*. Prentice Hall.
11. Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), 888-905, dostupan na: <https://ieeexplore.ieee.org/document/868688>
12. Serra, J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press.
13. Morphological Transformations, dostupan na: <https://docs.opencv.org/3.4/d9/d61/tutorial_py_morphological_ops.html>
14. Heijmans, H. J. A. M. (1994). *Morphological Image Operators*. Academic Press.